# Response to Reviewer 1 - 2025-28

## I. REVIEWER SUMMARY

In this manuscript, the authors present an alternative way to process "on-the-fly" (online) model output data, with the motivation being reduced data/memory usage and potential ease for users. Ultimately, this is an interesting manuscript, albeit weird. I think it can be made worthy of publication upon a revision. In general, the authors seem to be factual, careful, and nuanced. I think most of the work is relatively high quality. There are some problems though (the first one potentially fatal for the manuscript).

## II. MAJOR COMMENTS:

1. My reading of the manuscript is that it is making use of an algorithm package (whose version is in the title no less), but upon examination of the code availability section, I don't think the authors actually used the package at all in their analysis. Instead, they "simulated" using the package. I find this quite odd. I wrote a comment about this at the end. I don't know if this is intentional or not, but this is pretty deceptive. I invite the authors to explain, and I will keep an open mind.

   The reviewer raises a valid point here. To explain, the source code of the One_Pass package was not originally introduced due to contractual obligations. As we still felt that the analysis of the algorithm's performance was worthwhile, the backbone of the package was extracted so as to produce the same results, just not using the easy user interface and functionality of the One_Pass. This however was not deemed appropriate and when we were granted permission to make the package open source, we did so but did not update the corresponding notebooks. We have now updated the notebooks accordingly, in order to present the same Figures but produced with the code of the One_Pass package. We do note however that in some of the notebooks relating to the t-digests we also directly use the underlying t-digest package for analysis. This is because the One_Pass package acts as a user interface, hiding the background computation and will simply return the final percentile or histogram (depending on the user request). We do not expose the inner workings of the t-digest to the user. In the paper analysis however, we also provide details on cluster sizes and weights,

information not required when examining the final percentile but interesting for analysis and obtainable via the t-digest package. There was never an intent to be deceptive here and we have now more thoroughly integrated the One_Pass into the manuscript and used it in the making of the Figures.

2. The authors didn't explain how this type of algorithm/package would be used with an actual climate model running. Maybe it would be used in a futuristic cloud-compute setup? (They do cite some cloud-computing works.) I think this manuscript would greatly benefit from a detailed example (ideally a workflow) of how this would be implemented in an end-to-end fashion.

   We agree that the manuscript would benefit from a description of the overall setup. Originally this was not included due to limitations of not allowing the package to be open source, so we kept the scope of the paper to the more theoretical implementation. After clearing the contractual limitations that originally lead to the exclusion of the source code of the One_Pass package, we have included a section 'Design and implementation into a workflow' (Sect 2.). Here we describe the design choices of the package in the context of the project that it was built for. We explain the basic configuration and how it is currently used in an HPC workflow via a workflow manager.

3. I think this manuscript would benefit from a comparison to the "online diagnostics" route. In the "online diagnostics" route, climate scientists (and developers) write functionalities to output specific items of interests, without needing too much data. Climate models already do online mean, min, max, etc. in their output streams — nothing here is novel. That could easily be extended to all sorts of statistics. In fact, it could be extended in a composable fashion to even more complex and variable algorithms. For example, consider the following: "I want to calculate the globally averaged temperature at cloud-top but with a threshold of 280 K". An algorithm can be written to identify cloud top, then finding the temperature there, considering the threshold, and then doing a horizontal reduction in a composable way. This algorithm can then be run after each time step in the model, and storing in memory the variable to be output or writing it out immediately. I think the authors could improve this manuscript if they compare and contrast their one-pass streamed approach to that of models natively doing more online calculations themselves (potentially on idling CPUs, if the models run on GPUs) instead of an external module/package.

While it is true that models often provide mechanisms to compute variable statistics during the simulation runs, the key here is that a model no longer needs to be tailored to the specific needs of a data consumer downstream. The increasing resolution of climate models introduces much complexity when trying to tailor the model outputs to the requirements of the data consumer. While it is common for models to produce online calculations of monthly means of certain variables, getting different modelling groups to align on specific statistics for downstream users (e.g. 99th percentile of wind speed) would be challenging, to say the least. In the new section we describe the design choices behind the One_Pass package, and that it needs to be able to integrate with a variety of GCMs. We also outline the concept of the ClimateDT and that one of the ideas is that data consumers may enter or leave the simulation data stream at any time, according to their needs, and do not enforce their needs onto the design of the model. Instead, with the One_Pass they are provided with a mechanism to select their fields of interest, as well as the aggregation method and frequency for those fields. This will make the necessary operations at a later stage downstream and will be performed by an independent process, making the whole ecosystem more scaleable. We hope the inclusion of the new section has given this more context and also addresses the minor comment 7 for for L35–42.

Minor comments

1. L2: nitpick: “as such” is kinda weird here in that it made me think running at high-res was a consequence of the preceding sentence, but it actually is more of a antecedent of the following sentence.

   We’ve removed this along with the first sentence in an attempt to condense and consolidate the abstract.

2. L5: nitpick: probably misplaced comma (I’d move it to after HPC instead)

   Fixed.

3. L7: the phrase “data streaming” is introduced like a known quantity that is specific climate science, but in fact, it is neither known nor usually associated with climate science. Maybe jargon? If I were you, I would consider removing “data stream;” (the phrase with the semicolon) and let the sentence stand without the disruption

Data streaming is now properly defined in the abstract and in the introduction. It is a key concept that requires explaining as without projects adopting this method, there is no need for one-pass algorithms or the package. We have made sure to better define all jargon used in the manuscript.

4. L9: one-pass may necessitate explanation here (so I would rephrase)

   We've better defined these terms in the abstract and introduction making sure not to introduce jargon without explanation.

5. L9: intelligent doesn't seem like the right word (maybe efficient?) — these techniques aren't learning things, or are there?

   You're right, they're not learning algorithms, we've removed this adjective in the rephrasing. Also removed in the introduction.

6. L17: I think "well within the acceptable bounds" is an underestimated. Imho and understanding, these algorithms basically recover accuracy to within an insurmountable precision limit, so basically as good as one would get anyway. I would rephrase to finish with a stronger point

   Rephrased to make this point.

7. L35–42: I think this may surprise an average climate scientist involved in model development and evaluation (such as myself). We always had the ability to access data before simulations are done and one could obviously write stuff in a continuous manner. I think what you want to highlight is that you're algorithmically calculating some interesting statistics that capture specific interests. That is not really novel, not unique, but it is interesting and worthy of publication. In other words, I don't see the "novel" method and "unprecedented" reduction of "meaningful" output. Can you please (carefully) elaborate?

   We understand the reviewers concern here and agree the paradigm that the One_Pass package has been built around has not been fully explained. Indeed, someone involved in model development can write and access the climate model data before the whole simulation is finished. Here we are referring to a more flexible set up where data consumers who are not part of the model development can still 'tap into' the data stream. This means someone who would traditionally access data via an online archive could embed into the data stream and

start working with the data without having to wait for the full model run to be finished and published online. To address your concern we have modified the introduction significantly to convey data consumers as being those who don't necessarily work in the model development and traditionally don't have access to model data as it's being produced. We have also included the new section 2 (already detailed above) which puts the development of the One_Pass into the context of the project, with the eventual aim being anyone with an impact model or similar can embed into the data stream.

8. L70: the last sentence here gives the wrong impression of what you're trying to say. Your manuscript is supposed to showcase those algorithms, so "requesting" readers to go follow some other package documentation may not be the best way to politely say, "We don't discuss the code implementation of specific algorithmic details related to each statistic, and we instead focus on showing their utility to climate analysis" . . .

   We have more thoroughly embedded the differences between the One_Pass package and the aim of the paper, which as you say, is looking at the algorithms utility. We have rephrased L70 and tried to make it clear that examples of the package can be found online while here we discuss accuracy.

9. L194: I'd elaborate on the different results part

   We have expanded this statement to provide some more details however when sampling other packages we did not find substantial differences between the implementations. Our decision was to use this particular implementation was also based on how often the package was being maintained.

10. L219: I think you can say that plainly at the outset without taking the reader on an unnecessary voyage. I am not sure if the added info (analysis) is informative otherwise. Feel free to disagree.

    We realize this paragraph was quite convoluted. We have re-worded to introduce the main point at the start of the paragraph and have tried to condense to make the message clearer.

11. L268: so, the NumPy-calculated one is the "reference" right? (Note, I would refer to NumPy like they refer to it in their papers/docs, uppercase N and P)

Yes, NumPy is the reference. We've made this explicit in the introduction and fixed the casing.

12. Section 5.3: If I were you, I would simply delete this section. Your concluding paragraph (around L385) basically shows that most of the preceding analysis/"results" should be taken with a giant grain of salt. I would simply avoid the distraction, maybe add something about how the analysis in 5.2 could done on precipitation with a meaningful cutoff and/or a different type of underlying assumption (distribution type)

    We don't agree that removing this section is a good idea. The aim of the paper is to show how well the package will represent different statistics for different climate variables that are of interest to the climate community. Accurately capturing extreme precipitation is essential and we want to provide a detailed analysis of how the package performs on this front. We have however made some modifications to the section, reducing the text to make it more conscience and updating the Figure. Now we show comparisons between the histograms for the t-digest method and the NumPy method which aid in the explanation of the differences for the 99th percentile estimate.

13. L431: But this doesn't make sense to me. As a scientist, what I care about is the stability of the value, not an arbitrarily defined metric of convergence. In Figure 6, $\sigma_n$ goes from roughly 0.35 at 200 samples to 0.15 at 600 samples, but it is within those "convergence" lines.

    We use the convergence rate `https://en.wikipedia.org/wiki/Rate_of_convergence` to define how long it takes for the underlying statistic summary to be representative of the statistic of whole data set. We are not claiming the statistic has reached a final or static value. Climate variables will exhibit temporal variability and long-term trends, and we do not assume that their statistical properties remain fixed over time. Rather, our focus is on when the variability in the estimate of $\sigma_n$ itself has sufficiently diminished — i.e., when its fluctuations due to sample size limitations become negligible compared to inherent data variability. This is to allow the statistic summaries to be used in methods such as quantile-quantile mapping for bias-adjustment of streamed data. We have modified the introduction of the section considerably to clarify this.

14. L432: Ok, but how does that impact your assertions earlier about all the savings and not

having to wait for model simulations to run a lot of steps? I guess you're saying, one kinda has to wait a lot of steps to get "accurate" stuff out of these one-pass algorithms? Like 200 steps or so?

The analysis conducted here doesn't impact earlier assertions about the use of one-pass methods. As with previous sections, here we are benefiting from the memory reduction of the one-pass and the ability to work with the data stream. Our point about timely access was meant to compare, for example, those with impact models who currently use data from online archives. In the overall aim of the ClimateDT, these types of models will be able to embed into the ClimateDT workflow and access directly the model output via this streaming paradigm. We hope that with the inclusion of the new section this has been made much clearer. If you don't agree, we would greatly appreciate more thoughts on this.

15. Section 6: Like Section 5.3, I think the convergence analysis is likely misleading, potentially counterproductive, and perhaps better left out. Your goal in this manuscript is to showcase something useful for users of this method. My assumption is that if people choose to use this type of one-pass algorithm, they would do so on relatively high-frequency data and they would understand the risk of under-sampling. Maybe I am just not getting what you're trying to do here? Could you motivate it better? Is it anything other than something like "don't take the mean of the first 5 time steps as the mean of the next 5000 time steps"??

    We have re-written a lot of the beginning of this section to better explain the motivation behind it. As we were designing the package for bias-adjustment for streamed data (another package not presented here), we came across this issue of converge of statistic summaries and feel that it is a useful discussion for others. In this section we hope to give users guidance in the other ways that summaries from one-pass algorithms can be used: such as in quantile-quantile mapping, and what should be considered in the process. This is now (we hope) much better explained in the introduction to this section.

16. L508: Thanks for the links. A few questions/suggestions: 1) might be good to list the GitHub links here as well? 2) is there a hosted version of the package docs somewhere? 3) how was the one-pass package used in making these figures? I didn't see any "import one pass" or "from one pass…" … looks like you reimplemented/"simulated" everything from scratch? If so, this means the package in the title of this manuscript wasn't used at all and

the whole manuscript is misleading. I also don't quite understand the numerics are returning round-off errors, where I would've expected exact answers (e.g., for the means)

Yes, we have included also the GitHub links along with the link to the hosted version of the package. The notebooks have now been reformatted using the package and we apologize for the confusion that they weren't originally included.

---